Evaluation of the perceived colour difference under different lighting for museum applications

Alice Plutino[1], Laura Grechi[1], Alessandro Rizzi[1]

[1] MIPS Lab, Dipartimento di Informatica, Università degli Studi di Milano, Milano, Italy.

Corresponding Author: Alice Plutino (alice.plutino@unimi.it)

ABSTRACT

The role of lighting in museums has a fundamental importance. Light influences the perception of colours and space in the collections on display, therefore, any type of lighting must be adequately analysed to confirm the suitability and undistorted colour rendering of the illuminated objects. A two-stage perceptual test was carried out for this purpose. Initially, the participants were asked to evaluate the illuminants according to criteria such as: the brilliance of the colours, the degree of pleasantness of the lighting and the degree of overall satisfaction of the setting. Subsequently, the efficiency of different illuminants for the identification of colour differences between two objects was tested. The results obtained were then compared with the most commonly used colour rendering and colour difference indicators in order to determine their potential and limits.

KEYWORDS Lighting, Colour Perception, CRI, Museum Lighting

RECEIVED 31 October 2019; **REVISED** 16 November 2019; **ACCEPTED** 22 November 2019

1. Introduction

In museums, the role of lighting is often linked to the creation of an immersive and evocative experience for the viewer when visiting the collections. In this context, although the regulations for lighting that does not cause damage to the illuminated objects and to the users are respected, often the lights are not subjected to an analysis that confirms their adequacy and ensures a correct color rendering of the objects on display (Feller 1968, Camuffo 2014).

The experiment we are presenting can be divided into two main parts: evaluation of light sources and evaluation of colour differences.

In recent years, several experiments have been carried out in order to evaluate the perception of the colour of museum objects under different light sources (Boissard & Fontoynont, 2009, Pinto et al. 2008, Pinto et al. 2006, Scuello et al. 2004), focusing primarily on colour temperature and the study of LEDs. Nevertheless, it is difficult to integrate the elaboration of the visual system in the calculation of the colour perception and often it is reduced to evaluating a source based only on emission and reflectance spectra.

For this reason, it was decided to have users evaluate three main characteristics of the colour rendering: the brilliance of the colours, the pleasantness in relation to the lighting and the degree of overall satisfaction of the set-up. The aim of the brillance assessment is to determine how much the illuminant promotes the colour rendering of an object on display. The term pleasantness refers to the illuminant and is designed to determine which light is most pleasing to the viewer, not considering, therefore, the color rendering of the object. Finally, with the term satisfaction, we ask you to evaluate the entire installation.

In addition to illuminants, the study of colour rendering in museums in recent years has also been the subject of many studies, aimed at determining which source guarantees better colour rendering and, if there are preferable sources, for the lighting of specific museum objects (Viénot et al. 2011, Nascimento & Masuda 2012, Nascimento & Masuda 2014, Vázquez et al. 2012). In this context, in the second part of the experiment, users were asked to compare the colours of two Colour Checkers (see section 2), to determine how much the illuminant influences the perception of colour differences.

The results of these two experiments were compared with the most common colour rendering and colour difference indexes. The aim of this study is to demonstrate that colour rendering is a complex concept in its definition and that the most common indexes are not enough effective to represent the perceptual variations of colour. The applicability of the results obtained has been evaluated in the light of a complex field of application such as that of the museum, where for an exhibition lighting must be considered in terms of conservation, exhibition and use. In addition, the limits of colorimetric measurements, which are not reliable in the presence of spatial arrangements and different lighting conditions, are to be highlighted.

2. Materials and Methods

For this experiment, the use of the Gretag Macbeth Colour Checker, in original and printed version (Figure 2a), using a printer LaserJet Pro 400 colour. Both Colour Checkers have been positioned within a Light Booth, model The Judge II, equipped with 4 different types of light sources: CIE D65, Cool White Fluorescent, U30TL84 and Illuminant A. Two LED lights have been added to the Light Booth: one 'cold' (LED 1, 5000K) and one 'warm' (LED 2, 2500K). The uniformity of the light diffusion was verified on the basis of the Light Booth for all the lights.

The emission spectra of each individual source were measured with a CL-500A spectrophotometer (spectra in Figure 1, measurement set-up in Figure 2b) which, during the measurements, was placed in the centre of the Light Booth on the same support as the Colour Checkers. From the emission spectra of the sources the following were calculated: CRI (Colour Rendering Index) (Oleari, 1998), TM-30 Fidelity Index and TM-30 Gamut Index (Society, 2018).

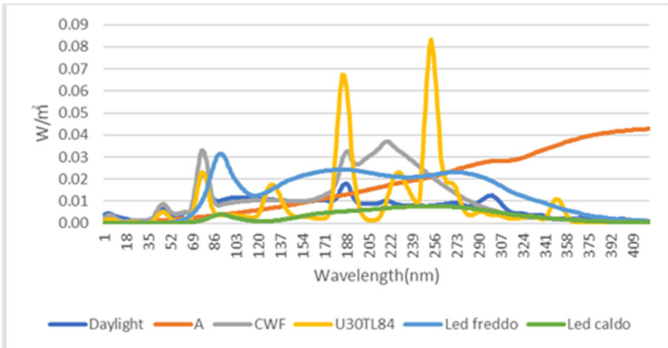


Fig. 1. Emission spectra of illuminants

The reflectance spectra of the patches of the two versions of the Colour Checker were measured using an Ocean Optics HR4000 spectrophotometer. The data thus obtained were then processed to obtain the colorimetric values in the L*a*b* colour space by simulating the illumination under the various light sources present in the Light Booth. Finally, the chromatic differences (values ΔE and $\Delta E00$) between the patches of the original Colour Checker and the corresponding ones of the printout were calculated.

The perceptual test was carried out on 20 people, none of whom had problems with colour perception.

The test, performed inside a dark room (Figure 3) without windows and without artificial lights on except for the Light Booth, was carried out in two parts:

a) Evaluation of the light sources: with the use of a guided table it was asked to indicate with an evaluation from 0 to 100 the brilliance of the colours, the degree of pleasantness referred to the light source and the degree of general satisfaction of the setting. The term brilliance was used because in the common language it is associated with the perceived brightness of the colours. It was decided not to use the term saturation or tone, as it is particularly specific and difficult for non-experts to understand.

This operation has been done alternating for each source, the original Colour Checker (McCamy et al. 1976, Pascale 2006) to the printed one so as to avoid that the judgment was subject to comparison between the two.

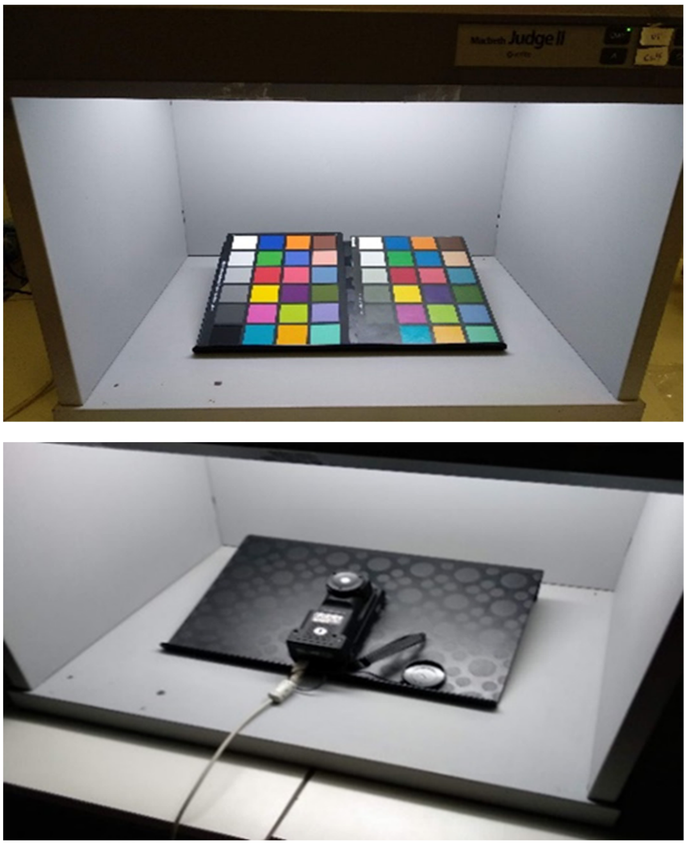


Fig. 2. (top) Original (left) and printed (right) colour checker inside the Light Booth. (bottom) Set-up of the CL-500A spectrophotometer during measurements.

b) Evaluation of colour differences: it was asked to indicate with an evaluation from 1 to 5, (where 1=identical patches and 5=absolutely different patches), the colour difference between 15 selected patches of the original colour checker and the corresponding printed patches (Figure 4) (McCamy et al. 1976, Pascale 2006). This subjective colour difference assessment was carried out under all Light Booth sources. To solve the numerical problem of estimating the magnitude between different colors, the "Neutrals" tab of the Munsell Book of Colours was used as a framework. For example, the magnitude between patches N9 (white) and patches N7/N8 (light grey) corresponds to a difference of 1, while patches N9 and N3/N2 (dark grey) have a difference of 3.

3. Results

3.1. *Evaluation of light sources*

The table below shows the CRI, colour temperature and TM-30 values of the Fidelity Index (Table 1). Figure 5 shows the results of the subjective evaluation of brilliance, pleasantness and satisfaction for each individual source.

	Daylight	A	CWF	U30TL84	LED 1	LED 2
CRI (Ra)	94	99	59	87	94	83
T colour(K)	7000	2600	3900	3000	4800	3000
Fidelity Index (Rf)	94	99	67	82	93	84

Tab. 1. CRI (Ra), T colour, Rf values of the various illuminants.

3.2. *Evaluation of colour differences*

The following table summarises the values ΔE and ΔE00 calculated from the reflectance spectra of each patch of the Colour Checker under all the sources present in the Light Booth (Figure 6). All calculated ΔE values have been compared with the average of the values assigned by users as shown in Figure 7, Figure 8, Figure 9, Figure 10 and Figure 11 and Figure 12.

4. Discussion

The results of the perceptual test showed some discrepancies between the calculated indexes and the perceived values in the evaluation of illuminants and colour differences.

3.2. *Evaluation of colour differences*

With regard to the assessments of brilliance, pleasantness and satisfaction, it is interesting to note that the scores awarded by users never reach the high scores awarded by the CRI index (Figure 5 and Table 1) (Fumagalli et al. 2013). The exception is the illuminant CWF of which the printed version of the Colour Checker has slightly lower perceptual values, but substantially similar, compared to those of the CRI, while the score assigned to the original version is higher.

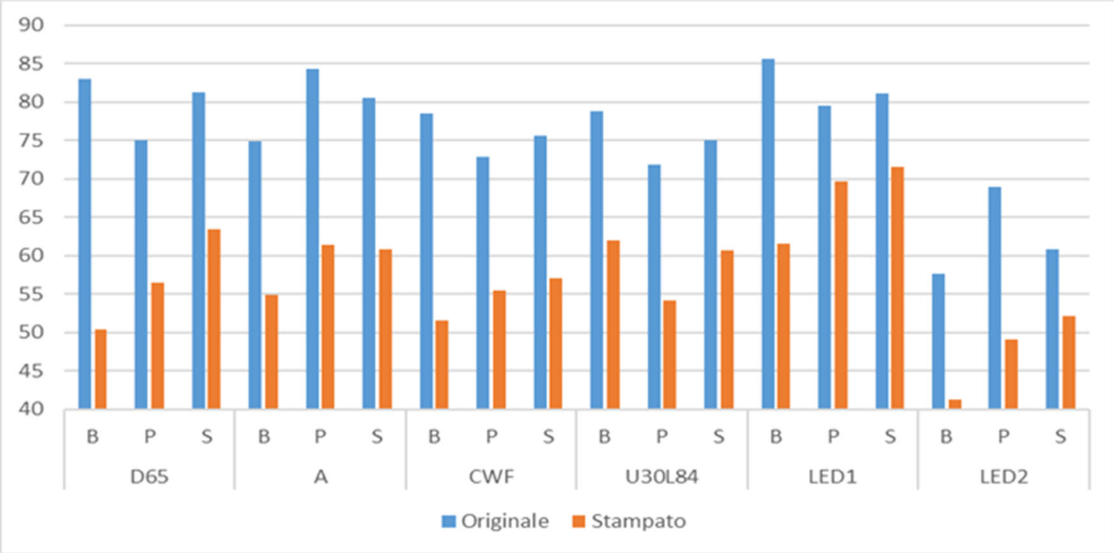
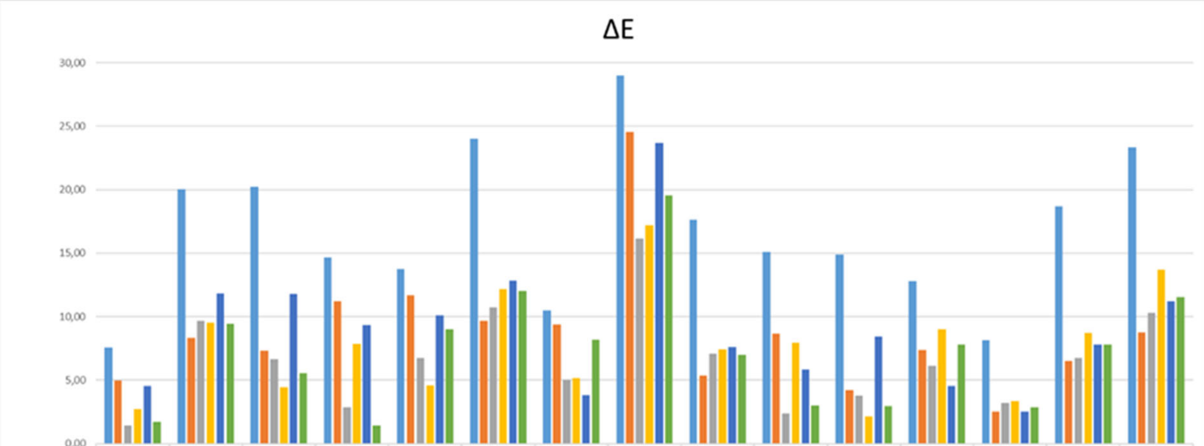
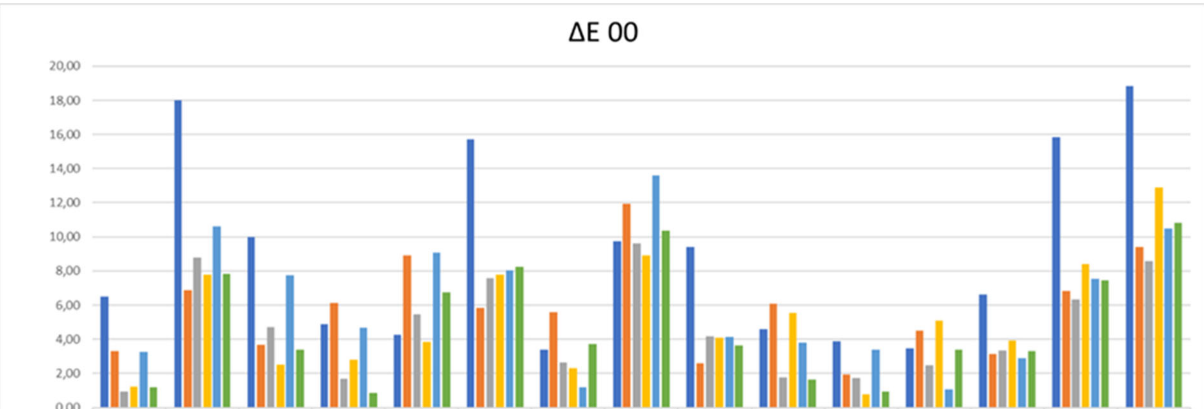


Fig. 5. User-assigned values in % for three parameters: B=Brilliance, P=Pleasantness, S=Satisfaction



	Dark Skin	Light Skin	Blue Sky	Foliage	Blue Flower	Bluish green	Orange	Purplish Blue	Moderate Red	Purple	Yellow Green	Orange Yellow	White 9.5	Neutral 6.5	Black 2
D65	7,56	20,00	20,22	14,67	13,77	24,00	10,49	28,98	17,65	15,07	14,89	12,81	8,12	18,67	23,34
A	4,99	8,36	7,34	11,22	11,68	9,68	9,38	24,52	5,35	8,67	4,20	7,36	2,51	6,49	8,75
CWF	1,43	9,68	6,65	2,87	6,77	10,74	5,02	16,12	7,09	2,39	3,80	6,11	3,21	6,75	10,28
U30TL84	2,75	9,53	4,47	7,85	4,59	12,15	5,16	17,18	7,43	7,94	2,13	8,99	3,36	8,72	13,69
LED1	4,53	11,85	11,78	9,32	10,12	12,86	3,83	23,69	7,64	5,83	8,45	4,52	2,53	7,81	11,19
LED2	1,71	9,45	5,54	1,41	9,02	12,00	8,17	19,55	6,99	3,00	2,95	7,82	2,87	7,80	11,57



	Dark Skin	Light Skin	Blue Sky	Foliage	Blue Flower	Bluish green	Orange	Purplish Blue	Moderate Red	Purple	Yellow Green	Orange Yellow	White 9.5	Neutral 6.5	Black 2
D65	6,50	18,01	9,99	4,86	4,25	15,71	3,38	9,76	9,39	4,60	3,89	3,46	6,61	15,83	18,85
A	3,30	6,86	3,68	6,13	8,91	5,84	5,60	11,92	2,58	6,10	1,95	4,50	3,14	6,85	9,40
CWF	0,91	8,79	4,71	1,70	5,45	7,57	2,63	9,61	4,16	1,77	1,71	2,46	3,36	6,33	8,57
U30TL84	1,23	7,77	2,51	2,82	3,82	7,80	2,29	8,93	4,08	5,56	0,75	5,10	3,93	8,42	12,89
LED1	3,24	10,62	7,74	4,67	9,08	8,04	1,17	13,58	4,14	3,82	3,40	1,07	2,89	7,55	10,47
LED2	1,20	7,83	3,40	0,85	6,75	8,24	3,73	10,36	3,64	1,65	0,95	3,38	3,30	7,46	10,80

Fig. 6. ΔE (top) and ΔE00 (bottom) values calculated under the different illuminants in the Light Booth

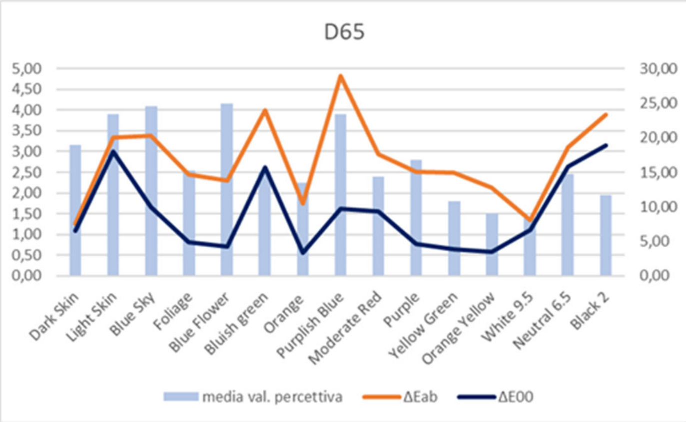


Fig. 7. Perceptual test values compared with the measured values of ΔE and ΔE00 for illuminant D65.

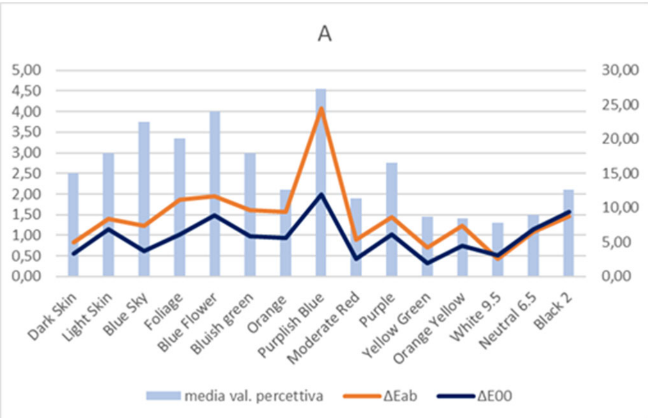


Fig. 8. Perceptual test values compared with the measured values of ΔE and ΔE00 for illuminant "A".

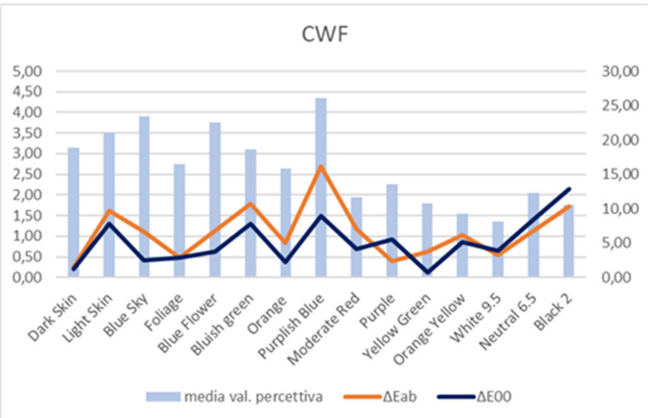


Fig. 9. Perceptual test values compared to measured values of ΔE and ΔE00 for illuminant "CWF".

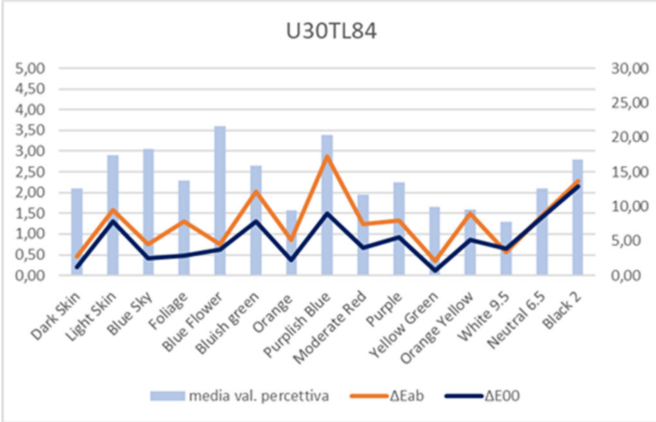


Fig. 10. Perceptual test values compared with the measured values of ΔE and ΔE00 for the illuminant "U30TL84".

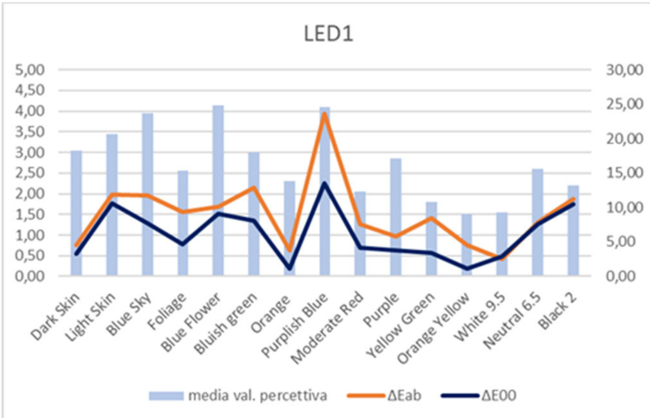


Fig. 11. Perceptual test values compared with the measured values of ΔE and ΔE00 for the LED illuminant1.

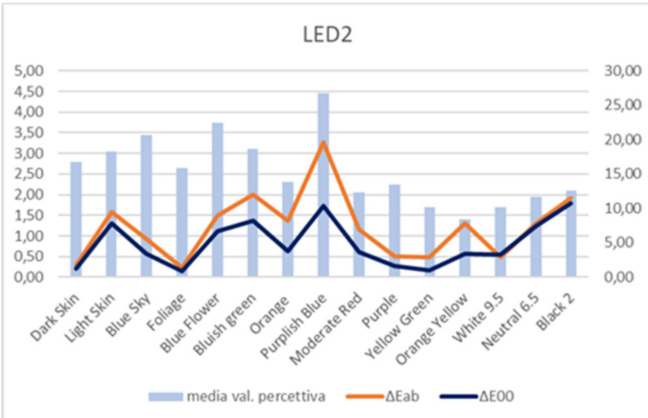


Fig. 12. Perceptual test values compared to measured values of ΔE and ΔE00 for LED illuminant2.

It is clear that the original version of the Colour Checker under all illuminants is much more appreciated by users, who always give it very high scores, indicating that the support and materials used for any coloured object have a fundamental importance for good colour rendering.

Considering the overall satisfaction value, the illuminants that received the highest scores from users were illuminant D65, A and LED1. This evaluation is consistent with the values of CRI and TM-30 Fidelity, as these measurements also assign higher values to the illuminants themselves. Note that the LED1 scored high according to these two indexes because, being of the new generation, the emission spectrum was modified by the production company in order to meet the CRI index standardized by the CIE commission. In addition, the values assigned by the TM-30 Fidelity Index do not differ particularly from the CRI.

Considering the values of brilliance and pleasantness, there are major discrepancies between the subjective judgment and the colour rendering indexes. According to the brightness values, in fact, the LED1 has the highest value, followed by A, CWL and U30TL84 for the original Colour Checker, while for the printed version the highest brightness is given by the U30TL84 and LED1. Considering the pleasant lighting value for the original, the best illuminant is A, followed by LED1, while the printed LED1 obtains the highest scores, followed by A.

These results cast doubt on the colour rendering indexes, which do not consider the perceptual component of colour processing by the subjects, as well as the colour adaptation and spatial context variables.

4.2. *Numbering and spacing*

With regard to the measurement of colour differences (ΔE and ΔE00), it can be seen that in general the values are always high (except in specific cases), due to the characteristics of the substrate and the dyes that have very different reflectance spectra. The purpose of the experiment is to highlight under what conditions the colorimetric and precise measurement of ΔE is not sufficient to determine the metering of two colours under different lighting conditions.

Considering the values of ΔE and ΔE00, it is evident that the measured values are significantly higher for the illuminant D65. In addition, when comparing the individual patches, higher colour differences are obtained for those with higher blue components (Bluish Green, Purplish Blue, Blue Sky, Blue Flower) (Brueckner et al. 2009). When comparing the illuminants, CWF and LED2 maintain the smallest difference values.

The comparison between the various illuminants does not reveal a general predominance for the subjective values compared to those calculated but depending on the patch and the type of lighting, the results assume very different values. Despite this, it is noticeable that patches with a strong blue component, assume values of high colour differences, while they generally assume lower values both for patches with a strong red component (Orange and Moderate Red) and for those in shades of grey (White 9.5, Neutral 6.5, Black 2), with a few exceptions.

Patches with a strong red component and patches in shades of grey obtained lower values both in the calculated ΔE and in comparison, to the scores assigned to all the others.

Considering illuminant D65 (Figure 7), the differences noted by the subjects are always greater than 2, with a peak of difference for the Purplish Blue that assumes values close to 5. Looking at the values for the illuminant A (Figure 8), there is a general decrease in the values of subjective difference, so the peak for patch 10 remains, but falling to a value of about 4.3. For White, Yellow Green and Moderate Red the difference values are close to 1.5 so they are considered perceptively very close to those of the original Colour Checker.

Considering the illuminant CWF (Figure 9), the values of Dark Skin and Purple are considered more similar to the original, while the colour difference of Purplish Blue returns to values of about 4.5. In this case the values of ΔE and ΔE00 are much lower than in the case of D65 and A.

As for U30TL84 (Figure 10) the difference for Yellow Green and Dark Skin is subjectively around 1.5 the values of ΔE and ΔE00 decrease for Purplish Blue but remains around 4.5 in subjective judgment. Finally, as far as LED1 (Figure 11) is concerned, the perceptual differences between the original and the printed one increase compared to the other illuminants and present a similar trend to D65, even if lowered by about 1 point. In LED2 (Figure 12) the subjective values are very similar to LED1, with some differences for some blue/green patches that have lower subjective values and some of the red-orange patches that have higher values.

From these evaluations and from the graphs we can see, therefore, a strong discrepancy between the ΔE and ΔE00 colorimetric values measured from the spectra to the perceptual values given by the evaluation of the subjects involved in the study, differences that are not given simply by a physical component of reflectance compared to an illuminant, but by a strong perceptual component. This is because both measurements ΔE and ΔE00 do not consider the spatial arrangement of colours in the evaluation and adaptation of the human visual system to the different illuminants.

5. Conclusion

In this study it is shown that the colour rendering indexes are not adequate to give an estimate of the perceptive rendering of an illuminating and to maintain metamerism. Moreover, it has been demonstrated that a colorimetric measure of local colour difference is not sufficient to represent the perceptual variations of colour with the variation of illuminants and in the presence of a spatial arrangement.

A further limitation to these measures lies in the fact that they do not consider the colour adaptation of the human eye, the surrounding illumination and the signal processing by our visual system. This aspect can be seen above all from the answers to the perceptual test on colour differences where it became clear that some users had the ability to recognize and preserve the variations between two coloured patches even under different illuminants. Moreover, in most cases there was no correspondence between the values acquired with the experimental measures and those assigned by the perceptual test.

For applications in the museum environment, provided that the regulations for lighting that does not damage the objects on display and the users are respected, it can be seen that as far as lighting is concerned, the supports and materials used in the enjoyment of the exhibition will always be of greater importance, so the object will always be of primary powerlessness. Moreover, it has been seen that LED lights far exceed the D65 standard in the creation of a satisfactory exhibition space and for a pleasant and apparently correct reproduction of colours.

Moreover, when evaluating sources for complex exhibition spaces, it must be taken into account that both the CRI and the Fidelity Index do not consider the spatial arrangement and the systems for adapting the eye to lighting, so if only one of these aspects is taken into account in the evaluation methods without considering the others, the comparison between the data and reality can never be considered satisfactory.

In conclusion, since many differences have emerged between the perceptual test and the values measured experimentally, both regarding CRI, ΔE and ΔE00, it is considered necessary to reconsider the methods and uses of both parameters to ensure that they are more reliable and consistent with reality.

Number the illustrations according to their sequence in the text. Submit each illustration as a separate file (indicate where the figure should be placed in the text). Ensure that each illustration has a caption. Supply captions separately, not attached to the figures. A caption should comprise a brief title (not on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

If visuals from other copyrighted publications are to be included in the manuscript, it is the author's responsibility to secure permission to publish these in the journal before submitting the manuscript for review.

6. Conflict of interest declaration

The authors declare that there is no conflict of interest with other people or organizations.

7. Funding source declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

12. Short biography of the author(s)

Alice Plutino - She is a PhD student in Computer science at the university of Milan. She obtained her master's degree in Conservation Science for Cultural Heritage at the same university. Now, she is focusing on colour correction, image enchantment and image quality assessments. She is author of different papers concerning innovative methods for film restoration and digital restoration assessments.

Laura Grechi - She is a graduate of the Master course in Conservation Science for Cultural Heritage at University of Milan. She obtained the master's degree in October 2019. During her curricular internship at the university she carried out experiments studying colour differences between human perception and colour rendering indexes.

Alessandro Rizzi - He is Full Professor at the Department of Computer Science at the University of Milan. From 1990 his research has been in the field of colour, digital imaging and vision. He is particularly focused on colour, visualisation, photography, HDR and on the perceptual issues related to digital imaging, interfaces and lighting. He is the head of the MIPS Lab at the Department of Computer Science and was one of the founders of the Italian Colour Group, Secretary of CIE Division 8, IS&T Fellow and Vice President. In 2015 he received the Davies medal from the Royal Photographic Society. He is co-chair of the IS&T Conference "Color Imaging: Displaying, Processing, Hardcopy and Applications", a member of several programme committees of conferences related to colour and digital imaging, and author of about 300 scientific works.

References

Boissard, S., & Fontoynont, M. (2009). Optimization of LED-based light blendings for object presentation. Color Res. Appl. 34, 310-320.

Brueckner, S., Bodrogi, P., & Khahhn, Q. T. (2009). Colour Rendering of new white LED light sources - visual tests. Lux Eurpoa, pp. 397-404.

Camuffo, D. (2014). Microclimate for cultural heritage - Conservation, restoration, and maintenance of indoor and outdoor monuments. Elsevier.

Feller, R. L. (1968). Control of deteriorating effects of light on museum objects: heating effects of illumination by incandescent light. . Museum News. Technical Supplement, 46(9), 33 - 47.

Fumagalli, S., Bonanomi, C., & Rizzi, A. (2013). An experiment on the color rendering of different light sources. Color Imaging XVIII: Displaying, Processing, Hardcopy and Applications, IS&T- SPIE Electronic Imaging, 3-7- February 2013, San Francisco (USA).

McCamy, C. S., Marcus, H., & Davidson, J. G. (1976). A Color Rendition Chart. Journal of Applied Photographic Engineering 11 (3), pp. 95-99.

Nascimento, S. M., & Masuda, O. (2012). Psychophysical optimization of lighting spectra for naturalness, preference, and chromatic diversity. JOSA A, 29(2), A144-A151.

Nascimento, S. M., & Masuda, O. (2014). Best lighting for visual appreciation of artistic paintings—experiments with real paintings and real illumination. JOSA A, 31(4), A214-A219.

Oleari, C. (1998). Misurare il colore. Hoepli.

Pascale, D. (2006). RGB Coordinates of the Macbeth ColorChecker.

Pinto, P. D., Linhares, J. M., & Nascimento, S. M. (2008). Correlated color temperature preferred by observers for illumination of artistic paintings. JOSA A, 25(3), 623-630.

Pinto, P. D., Linhares, J. M., Carvalhal, J. A., & Nascimento, S. M. (2006). Psychophysical estimation of the best illumination for appreciation of Renaissance paintings. Visual Neuroscience, 23, 669-674.

Scuello, M., Abramov, I., Gordon, J., & Weintraub, S. (2004). Museum lighting: Optimizing the illuminant. Color Res. Appl., 29, 121-127.

Society, I. E. (2018). IES Methods for Evaluating Light Source Color Rendition. ISBN-13: 978-87995-379-9.

Vázquez, D., De Luna, J. M., Alvarez, A., Sánchez, A., & Sedano, U. (2012). Study of chromatic variations between metameres by varying the lighting in the painting" Boy in a turban holding a nosegay. Optical Systems Design, International Society for Optics and Photonics, 8550.

Viénot, F., Coron, G., & Lavédrine, B. (2011). LEDs as a tool to enhance faded colours of museums artefacts. Journal of Cultural Heritage. 12(4), 431-440.